A review on applications of genetic algorithm for artificial neural network

Ali Saud Al Tobi, Maamer; Bevan, Geraint; Wallace, Peter; Harrison, David; Ramachandran, K. P.

Published in:
International Journal of Advance Computational Engineering and Networking

Publication date:
2016

Document Version
Author accepted manuscript

A review on applications of genetic algorithm for artificial neural network

*1Maamar Ali Saud Al Tobi. 2Geraint Bevan, 2Peter Wallace, 2David Harrison, 1K.P.Ramachandran

1Caledonian College of Engineering
P.O.Box: 2322, CPO Seeb 111
Sultanate of Oman
E-mail: altubi555@gmail.com, ramkp@caledonian.edu.om
*Corresponding Author

2Glasgow Caledonian University
Cowcaddens Road, Glasgow, G4 0BA
Scotland, UK
E-mail: geraint.bevan@gcu.ac.uk, D.Harrison@gcu.ac.uk, p.wallace@gcu.ac.uk

Abstract

With the fast growth of industrial machinery systems and considering the essential demand for the best machinery condition monitoring systems; artificial neural network (ANN) can be used as an automatic classifier for the fault of machines to avoid the human interpretation errors. Multilayer Perceptron neural network (MLP) is one of the popular ANN type due to its simplicity and efficiency. On other hand, Genetic Algorithm (GA) can be combined with MLP-ANN for the classification optimizing through methods based selection and training. This paper reviews and discusses the applications of GA with ANN and the future scope of applying GA for the training of ANN based machinery fault diagnosis in order to fill the gaps of traditional Back Propagation algorithm (BP).

Keywords: Artificial Neural Network, Genetic Algorithm, Multilayer Perceptron, Back Propagation, Training Algorithm.

1- Introduction

Machinery condition monitoring has been developed through different stages starting from the conventional methods (non-automatic) using time domain analysis [**1**] and frequency domain analysis, where methods such as the Fast Fourier Transform (FFT) are applied [**2**, **3**]. More recently, ANN with its many types have been applied as automatic fault diagnosis systems for the different rotating machineries and components such as Back Propagation-Artificial Neural Network (BP-ANN) or Multilayer Perceptron (MLP) [**4-14**], Radial Basis Function (RBF) [**4**, **6**, **9**], Probabilistic Neural Network (PNN) [**4**, **6**], to Support Vector Machine (SVM) [**8**, **10**- **12**, **15**, **14**].

Combining GA with MLP-ANN has been applied on machinery fault diagnosis for the selection of input features [**14**, **6**, **13**] and for the selection of the number of neurons in the hidden layer [**4**, **14**, **6**], where the most appropriate features and number of neurons have to be selected using GA. Optimizing the used number of features and neurons can improve the speed of training and the classification performance. GA has been applied to fault diagnosis of rotating machinery for feature and neurons selection. However, it has not been used for direct ANN classifier training, and has not been at all used for some machines like centrifugal pumps. Some studies on non-machinery based classification have introduced GA as a training algorithm where GA was used as a direct training algorithm [**16**-**18**] and combined with other training algorithms which is mostly back propagation (BP) [**19**- **23**].

This paper aims to discuss the current application of GA with ANN based machinery fault diagnosis which are limited into methods like the selection of number of features and neurons, whereas, GA has wider scale of applications like training method as some relevant literature on non-machinery would be reviewed. It is divided into six parts, including the introduction. Section 2 presents a brief review on ANN and MLP-ANN considering the advantages and disadvantages. Section 3 briefly reviews BP. Section 4 is subdivided into two sections: an introduction that gives a brief overview of GA; and a review of Genetic Algorithm for ANN based rotating machinery fault diagnosis. Section 5 reviews training methods based GA and GABP-ANN. Finally, a conclusion with remarks and future research scope is given in section 7.

2- Artificial Intelligence Neural Network

This section discusses briefly the definition of ANN and particularly Multilayer Perceptron (MLP) as one of the popular AI type.

Automatic fault detection methods make use of Artificial Intelligence (AI) which seeks to replicate mental capabilities with the support of computational systems [**24**]. Artificial neural network (ANN) was first introduced by McCulloch and Pitts in 1943 [**25**].

MLP consists of three layers, namely, input, hidden, and output layer of neurons. There may be several hidden layers between the input and output layers. The number of neurons in each section affects the generalization ability of the system, while the number of neurons and hidden layers affects the efficiency of the system. With larger number, there is a possibility of over-fitting the training data and weak generalization of new data. Therefore, some methods might be used to select the proper number of hidden layers and neurons such as Genetic Algorithm **[6**]. The output layer can be more than one layer based on the required fault classifications. Each hidden layer has a number of neurons; the role of each is to calculate the weighted sum of its inputs and apply the sum as the input of an activation function that is usually a sigmoid function. Back Propagation algorithm has been widely used in training of MLP. Figure 2-8 depicts the basic structure of a MLP network [**26**].

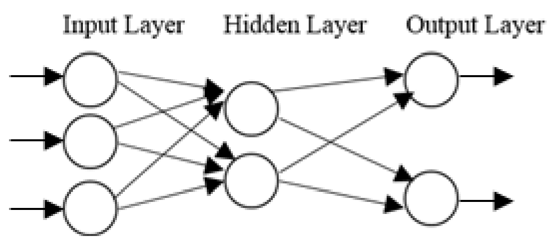

Figure 2-8: The basic structure of MLP network [26]

Comparative studies have demonstrated the efficiency of MLP over other ANN types [**4**, **6**] while considering important factors that affect efficiency, such as the number of hidden layers and neurons [**6**]. However, a drawback of MLP is that it is slow in training and needs longer time than other methods [**4**, **6**, **27**, **9**]; but such weakness can be minimized by reducing the number of input features [**5**, **7**]. Applying an alternative training method like GA rather than conventional BP, it can enhance training efficiency. It is also possible to combine GA and BP as a hybrid training method which gives better results. The details will be discussed in next sections.

3- Back-Propagation Learning Algorithm

Back Propagation (BP) is a training algorithm that was developed by Rumelhart in 1986 [**28]** and mostly its popularity has been built as training method for MLP. The basic training principle of BP is based on the gradient decent method that it works to adjust and modify the weights by minimizing Mean Square Error (MSE). The training process of BP can be started by multiplying the input vectors with the weights as the biases and weights are summed in order to calculate the actual outputs. Desired outputs which have to be determined and then compared with the actual outputs with continues evaluation and weights modification till the process approach the desired MSE value where then the training has to stop. MSE is also known as training or network error and represented mathematically as [**26**]:

$E = \sum_k \frac{1}{2}(t_k - y_k)^2$ **(1)**

Where t_k is the desired output in the output layer, and y_k is the actual output in the output layer.
Despite the popularity of BP as a training algorithm for MLP, it has some disadvantages that it is stuck with local minima during training [**19-21**] and it is slow that needs high number of iterations while training [**26**]. More discussion on the weaknesses and current alternatives of BP is given in section 4.

4- Genetic Algorithm

Introduction

Genetic algorithm (GA) was introduced by John Holland in 1975 [**29**]. It is based on the concept of a Darwinian-type fitness for survival that it is used to produce better individuals for the desired problem, as different possible solutions compete and match with each other. It is essentially a form of optimization, which can be applied to complex functions. GA has a similarity with chromosomes in that individual terms are represented by means of a linear string [**26**]. The basic concept of GA processes can be illustrated in Figure 2-12. GA starts its process by initiating individual populations which is known as chromosomes where they then would be computed and evaluated individually based on fitness and then they would be ranked according to the higher fitness where after that selection would be based on the top survival individuals (higher fitness). GA has two main operators, namely, crossover and mutation, and they would operate to reproduce new generation of individuals (chromosomes) and then would be sent to the first step of the process as the improper individuals are replaced with the new and good ones [**26**].

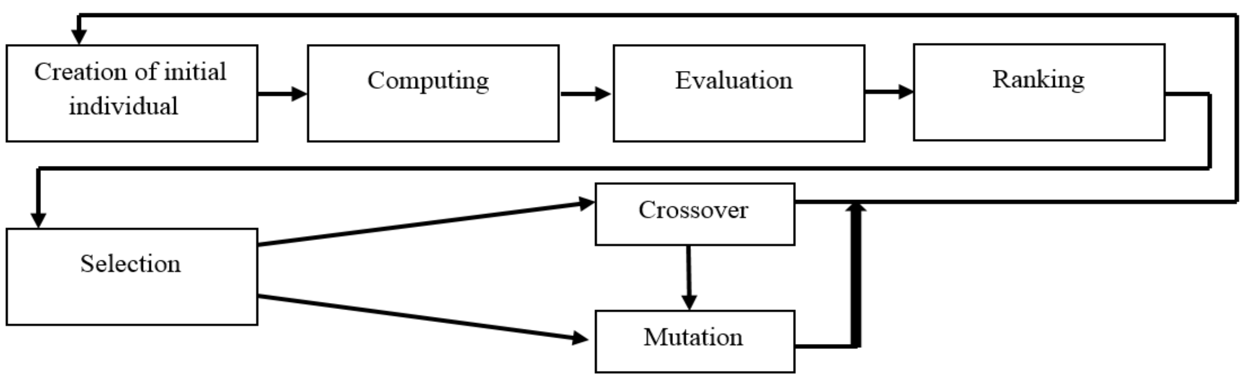

Figure 2-12: Main processes and operations of GA.

4.1. Genetic Algorithm for ANN based rotating machinery fault diagnosis

Samanta (2004) [**14**] proposed GA for the automatic selection of input features and number of neurons in the hidden layer. Two artificial intelligent classifiers were tested to diagnose gear faults, namely, Support Vector Machine (SVM) and ANN based Multilayer perceptron (MLP). This study was implemented with GA and without GA. Results showed that without GA, SVM outperformed MLP. However, using GA, the results were improved for both classifiers and classification success rate of both systems was 100%.
Samanta et al. (2006) **[6]** presented GA combined with three types of ANN classifiers; MLP, Radial Basis Function (RBF) and Probabilistic Neural Network (PNN) for fault diagnosis of bearing fault. GA was proposed for the best selection of input features and number of neurons in hidden layer. The input features were normalized where each row was divided by its absolute maximum value, and this method was applied to speed up and optimize process of training. It has been shown that using GA-based selection for feature selection, classification success rates were improved, as shown in Tables 1 and 2. GA is used to select only the best or appropriate features, which should thus improve the classification rate. In this study, 6 features yielded the best success rate for PNN and MLP, and 8 features gave the best performance with RBF.

Table 1: Without GA-based selection (All 45 features for each) [6].

RBF	PNN	MLP
83.33%	95.83%	85.06%

Table 2: With GA-based selection [6].

Number of Features	RBF	PNN	MLP
3	87.50%	96.53%	99.31%
6	95.14%	100%	100%
8	99.31%		

Al-Raheem and Abdul-Karem (2010) [**4]** proposed GA for the automatic selection of number of neurons in the hidden layer. Three types of ANN classifiers were proposed; MLP, RBF and PNN for the fault diagnosis of rolling element bearing. Input features were normalized between [0, 1] in order to be optimized. The results presented 100%, 97.5% and 72.1% as success rates for MLP, PNN and RBF respectively.
Yang et al. (2011) [**13]** applied GA for the best selection of input features where to be forwarded to MLP based back propagation algorithm (BP). The result has shown a classification success rate of 95% with GA-based feature selection as only the most appropriate features were selected using GA. This study presented the results based on combining GA along with ANN classifier only and comparison with the without GA-based feature selection was not presented.

From the above reviewed literature, it is observed that ANN based GA has been applied in rotating machinery fault classification, but within a limited scale wherein GA based training has not yet been used for rotating machinery. Section 5 reviews and discusses application of ANN-GA based training on other literature.

5- GA and GABP-ANN training methods

There are some reasons derived some researchers to think about an alternative for the traditional learning algorithm BP that the disadvantage of BP is being faster in facing the local minima during the training process which gets it stuck and not able to find the useful local information. GA has come as alternative where it has a global search ability with considering its slow comparing BP, whereas BP provides better local searching ability than GA [**19**]. Therefore, some researchers have preferred to combine GA with BP as a hybrid training method in order make each algorithm works to cover the other weaknesses [**19**, **20**, **21**]. Unfortunately, it has not seen any of such hybrid training method or even using GA instead of BP for rotating machinery based fault diagnosis including centrifugal pump. The following review literature are for different studies that are not in the area of machinery condition monitoring.

Guptaa et al. (1999) [**17]** compared GA as a training algorithm for MLP with the traditional BP using data of chaotic systems. The obtained results showed that GA provided better performance than BP, it was easier in used, and had more efficiency. Mcinemey and Dhawan (1993) [**20]** proposed a hybrid training method where GA and BP were combined to train MLP. Their proposed method was to start using BP to train the weights of MLP until reaching the local minima and then allowed to GA to continue the training after the weights were encoded into chromosomes. GA used the local minima instructions that were identified by BP to find the global optimum solutions. This method proved its efficiency as the drawback of each algorithm was overcome using the hybrid method. Kitano (1990) [**21**] proposed the same method before Mcinemey and Dhawan (1993) [**20**] but with reverse process that he first applied GA to approach the global optimum as GA gets stuck at this stage and to overcome this stuck, BP was allowed to continue training using the global instructions to find the local minima. It can remarked that such hybrid methods were so effective and the challenges of training were solved.

Siddique and Tokhi (2001) [**16**] presented a study to compare GA with BP for training MLP. They approached that GA was better than BP as it was able to do larger and global search.

Ding et al. (2011) [**19**] reported that the disadvantage of BP is that easily get stuck at the local minima and has a poor rate of convergence. This study proposed a hybrid training method that combined GA with BP in order that GA would optimize BP to overcome its weakness of getting stuck at the local minima. This method proposed that GA used to train the connection weights where these weights can be coded into chromosomes using two methods: binary encoding that each weight is represented by fixed length 0,1 string, and using real encoding that each weight is represented by a real number. This study used the real encoding to convert the weights into chromosomes. GA used to find better weights space, then BP used to complete training by searching locally in this space. Results showed that using hybrid training method (GABP) was provided better accuracy rate than using GA and BP separately as shown in Table 2-9. However, consumed time was longer with GABP.

Table 2-9: The comparison of the three algorithms [19].

Algorithm	Time (s)	Iterations	Accuracy rate (%)	Successful
BP	13.9	6577	93	9
GA	17.5	2067	92	24
GABP	16.6	6819	95	23

Mohmoudabadi et al. (2009) [**30**] proposed hybrid training method using GA and BP for MLP. This study was conducted for grade estimation. At the beginning, GA was applied to finding the optimal initial weights for network and then the network was trained till it reached its minimal MSE error using BP. This hybrid method was proposed to avoid the sensitivity of initial weights value as they drive the network to face local minima, and hence, this method had solved this problem using GA to find the optimal weights and minimize the sensitivity.

Güler et al. (2005) [**22**] proposed GA based training algorithm for MLP for lung sounds classification. GA was used to optimize BP by re-arrange the network weights, selection of number of inputs, and number of hidden neurons. It was also used to optimize the parameters of BP; the value of learning rate, and the momentum rate. Such values and parameters were encoded into chromosomes using two methods for encoding, namely, binary and real as shown in Figure 2-13.

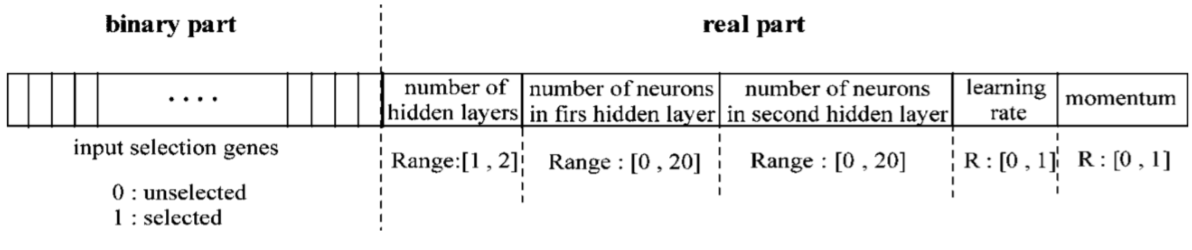

Figure 2-13: GANN chromosome structure using binary and real encoding [22].

Pallavi and Vaithiyanathan (2013) [**23]** combined GA with BP to train MLP for weather forecasting. The proposed method is illustrated in Figure 2-14 where GA was used to find the optimal value of weights as it was used for weights training. The weights were encoded into chromosomes where they then were calculated and evaluated using GA and calculation was based on MSE error. The chromosomes with higher fitness were selected and new generation of chromosomes of optimal weights value were produced using GA operators of crossover and mutation. Finally, the optimized weights were trained using BP. This method was effective.

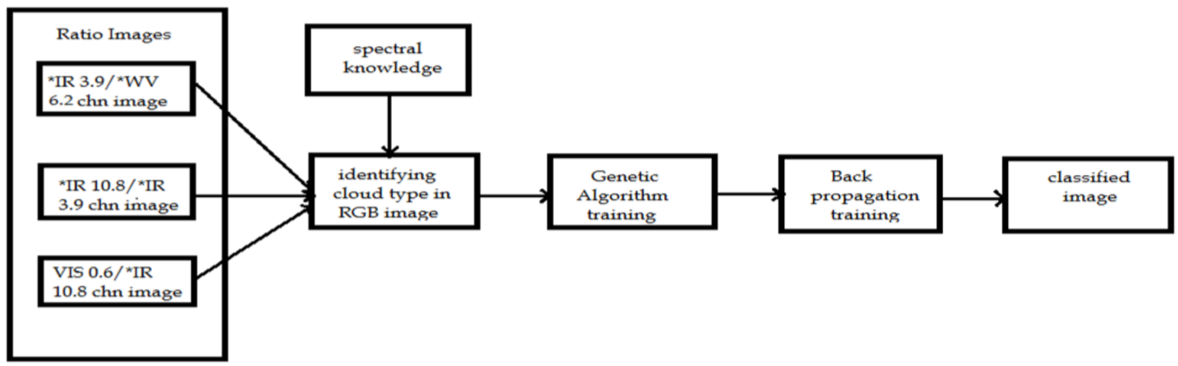

Figure 2-14: The proposed hybrid training method for MLP [23].

However, with another study by Joy (2011) [**18]** GA was compared with BP for training of MLP for pattern recognition problems and he found that BP outperformed GA.

6- Conclusion

This paper has reviewed the applications of GA for ANN based machinery fault diagnosis which have been observed to be limited into methods like the selection of number of features and neurons, whereas, GA has a wider scale of applications like training method as has been remarked through some relevant literature on non-machinery literature. Therefore, selecting of the appropriate training algorithms is important to speed up the process of training to obtain the best classification accuracy. GA has been applied successfully for non-machinery systems as a training algorithm, but, it has been noted to be with better performance when it is combined with BP algorithm as a hybrid training method. Such hybrid training method has not yet tested for any rotating machinery work.

References

1. AL-TUBI, M. A. S. & AL-RAHEEM, K. F. (2010) Rolling element bearing faults detection, a time domain analysis. *Caledonian Journal of Engineering,* 6.
2. AL-TUBI, M. A. S., AL-RAHEEM, K. F. & ABDUL-KAREM, W. (2012) Rolling element bearing element faults detection, power spectrum and envelope analysis. *International conference on applications and design in mechanical engineering.* Penang, Malaysia.
3. MEHALA, N. & DAHIYA, R. (2008) A Comparative Study of FFT, STFT and Wavelet Techniques for Induction Machine Fault Diagnostic Analysis. *International conference of computational intelligence, man-machine systems cybernetics (CIMMACS '08).* India.
4. AL-RAHEEM, K. F. & ABDUL-KAREM, W. (2010) Rolling bearing fault diagnostics using artificial neural networks based on laplace wavelet analysis. *International journal of Engineering, Science and Technology,* 2**,** 278-290.
5. AL-RAHEEM, K. F., ROY, A., RAMACHANDRAN, K. P., HARRISON, D. K. & GRAINGER, S. (2008) Application of the Laplace- Wavelet Combined With ANN for Rolling Bearing Fault Diagnosis. *Journal of Vibration and Acoustics,* 130.

6. SAMANTA, B., AL-BALUSHI, K. R. & S.A.AL-ARAIMI (2006) Artificial Neural Networks and Genetic Algorithm for Bearing Fault Detection. *Soft Computing,* 10**,** 264-271.

7. ZHANG, Z., WANG, Y. & WANG, K. (2013) Intelligent fault diagnosis and prognosis approach for rotating machinery integrating wavelet transform, principal component analysis, and artificial neural networks. *The international journal of advanced manufacturing technology,* 68**,** 763-773.
8. JEDLINSKI, Ł. & JONAK, J. (2015) Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform. *Applied Soft Computing,* 30**,** 636–641.
9. WANG, D. Y., ZHANG, W. Z. & ZHANG, J. G. (2010) Fault Bearing Identification Based on Wavelet Packet Transform Technique and Artificial Neural Network. *International conference on system science, engineering design and manufacturing informatization.* Yichang, IEEE.
10. YANJUN, L., FANLONG, M. & YIBO, L. (2009) Research on Rub impact Fault Diagnosis Method of Rotating Machinery Based on Wavelet Packet and Support Vector Machine *International Conference on Measuring Technology and Mechatronics Automation.* Changchun, IEEE.
11. SUI, W.-T. & ZHANG, D. (2009) Rolling Element Bearings Fault Classification Based on SVM and Feature Evaluation. *Proceedings of the Eighth International Conference on Machine Learning and Cybernetics.* Baoding, IEEE.
12. LIU, Y. & LIU, T. (2010) Rotating Machinery Fault Diagnosis Based on Support Vector Machine *International Conference on Intelligent Computing and Cognitive Informatics.* Kuala Lumpur, IEEE.
13. YANG, Z., HOI, W. I. & ZHONG, J. (2011) Gearbox Fault Diagnosis based on Artificial Neural Network and Genetic Algorithms. *International Conference on System Science and Engineering.* Macau, China.

14. SAMANTA, B. (2004) Gear fault detection using artificial neural networks and support vector machines with genetic algorithms. *Mechanical Systems and Signal Processing* 18**,** 625–644.

15. ZHONG, J., YANG, Z. & WONG, S. F. (2010) Machine Condition Monitoring and Fault Diagnosis based on Support Vector Machine. *International conference of Industrial Engineering Management* Macao, IEEE.
16. SIDDIQUE, M. N. H. & TOKHI, M. O. (2001) Training Neural Networks: Backpropagation vs Genetic Algorithms *International joint conference on neural network* Washington, DC, IEEE.
17. GUPTAA, J. N. D. & SEXTON, R. S. (1999) Comparing backpropagation with a genetic algorithm for neural network training. *The international journal of management science* 27**,** 679-684.
18. JOY, C. U. (2011) Comparing the Performance of Backpropagation Algorithm and Genetic Algorithms in Pattern Recognition Problems. *International Journal of Computer Information Systems,* 2.
19. DING, S., SU, C. & YU, J. (2011) An optimizing BP neural network algorithm based on genetic algorithm. *Artificial intelligence review,* 36.
20. MCINEMEY, M. & DHAWAN, A. P. (1993) Use of Genetic Algorithms with Back Propagation in Training of Feed-Forward Neural Networks. *International conference on neural networks.* San Francisco, CA, IEEE.
21. KITANO, H. (1990) Empirical studies on the speed of convergence of neural network training using genetic algorithm. *Machine Learning*.
22. GU¨LER, I., HU¨SEYINPOLAT & ERGU¨N, U. M. (2005) Combining Neural Network and Genetic Algorithm for Prediction of Lung Sounds. *Journal of medical systems,* 29.
23. PALLAVI, V. P. & VAITHIYANATHAN, V. (2013) Combined Artificial Neural Network and Genetic Algorithm for Cloud Classification *International Journal of Engineering and Technology (IJET),* 5.
24. CHARNIAK, E. & MCDERMOTT, D. (2000) *Introduction to artificial intelligence*, addison wesley Longman Inc.

25. MCCULLOCH, W. S. & PITTS, W. (1943) A logical calculus of the ideas immanent in nervous activity. *BULLETIN OF MATHEMATICAL BIOPHYSICS,* 5**,** 115-133.

26. ROOIJ, A. J. F. V., JAIN, L. C. & JOHNSON, R. P. (1996) *Neural Network Training Using Genetic Algorithm,* Singapore, World Scientific Publishing Co. Pte. Ltd.

27. LIU, L.-S., YANG, Y.-H., LI, Z.-Y. & YU, W. (2011) Condition Monitoring for Helicopter Main Gearbox Based on Wavelet Packet Transform and Wavelet Neural Network. *International conference on quality, reliability, maintenance, and safety engineering.* Xi'an, IEEE.
28. RUMELHART, D. E., HINTON, G. E. & WILLIAM, R. J. (1986) Learning representations by back-propagating errors. *Nature,* 323.
29. HOLLAND, J. H. (1975) *Adaptive in natural and artificial systems,* Ann Arbor, The university of Michigan press.
30. MAHMOUDABADI, H., IZADI, M. & MENHAJ, M. B. (2009) A hybrid method for grade estimation using genetic algorithm and neural networks. *Computational Geoscience,* 13**,** 91-10.